

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/538,471  
Source: PCT  
Date Processed by STIC: 06-28-2005

# ***ENTERED***

**CRF Errors Edited by the STIC Systems  
Branch**

Serial Number: 101538,471

CRF Edit Date: 06-28-2005  
Edited by: YF

\_\_\_ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

\_\_\_ Corrected the SEQ ID NO. Sequence numbers edited were:

\_\_\_\_\_

\_\_\_ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

\_\_\_\_\_

\_\_\_ Deleted: \_\_\_ invalid beginning/end-of-file text ; \_\_\_ page numbers

\_\_\_ Inserted mandatory headings/numeric identifiers, specifically:

\_\_\_\_\_

\_\_\_ Moved responses to same line as heading/numeric identifier, specifically:

\_\_\_\_\_

✓  
\_\_\_ Other:  
MOVED OVERLAPED <221> to <220> to  
the next line, as <220> does not require  
any response.

Raw Sequence Listing before editing,  
for reference only



PCT

## RAW SEQUENCE LISTING

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:16

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

```

3 <110> APPLICANT: Balakireva, Larissa
5 <120> TITLE OF INVENTION: MOLECULES INHIBITING HEPATITIS C VIRUS PROTEIN SYNTHESIS AND
METHOD FOR
6     SCREENING SAME
8 <130> FILE REFERENCE: 1759.200
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/538,471
C--> 11 <141> CURRENT FILING DATE: 2005-06-03
11 <150> PRIOR APPLICATION NUMBER: PCT/FR03/03675
12 <151> PRIOR FILING DATE: 2003-12-11
14 <150> PRIOR APPLICATION NUMBER: FR0215718
15 <151> PRIOR FILING DATE: 2002-12-12
17 <160> NUMBER OF SEQ ID NOS: 16
19 <170> SOFTWARE: PatentIn version 3.1
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 326
23 <212> TYPE: DNA
24 <213> ORGANISM: Artificial Sequence
26 <220> FEATURE:
W--> 27 <221> NAME/KEY: HCV
28 <222> LOCATION: 40..372
29 <223> OTHER INFORMATION: corresponds to IRES sequence of HCV
31 <400> SEQUENCE: 1
33 ctccccctgtg aagaactact gtcttcacgc agaaagcgtc tagccatggc gttagtatga      60
35 gtgtcgtgca gcctccagga cccccctcc cgggagagcc atagtggctc gcggaaccgg      120
37 tgagtacacc ggaattgcca ggatgaccgg gtcctttctt ggatcaacct gctcaatgcc      180
39 tggagatttg ggcgtgcccc cgcgagactg ctagccgagt agtggtgggt cgcgaaaggg      240
41 cttgtggtac tgcctgatag ggtgcttgcg agtgccccgg gaggtctcgt agaccgtgca      300
43 tcatgagcac aaatcctaaa gaaaaa      326
46 <210> SEQ ID NO: 2
47 <211> LENGTH: 80
48 <212> TYPE: DNA
49 <213> ORGANISM: Artificial Sequence
51 <220> FEATURE:
W--> 52 <221> NAME/KEY: HCV
53 <222> LOCATION: 40..119
54 <223> OTHER INFORMATION: corresponds to a portion (region II) of HCV IRES sequence
56 <400> SEQUENCE: 2
58 ctccccctgtg aggaactact gtcttcacgc agaaagcgtc tagccatggc gttagtatga      60
60 gtgttggtgca gcctccagga      80
63 <210> SEQ ID NO: 3
64 <211> LENGTH: 37
65 <212> TYPE: DNA
66 <213> ORGANISM: Artificial Sequence
68 <220> FEATURE:

```

## RAW SEQUENCE LISTING

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:16

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

W--> 69 <221> NAME/KEY: HCV  
 70 <222> LOCATION: 56..92  
 71 <223> OTHER INFORMATION: corresponds to a portion (consensus sequence) of HCV IRES  
 sequence

73 <400> SEQUENCE: 3  
 75 tactgtcttc acgcagaaag cgtctagcca tggcggt 37  
 78 <210> SEQ ID NO: 4  
 79 <211> LENGTH: 814  
 80 <212> TYPE: PRT  
 81 <213> ORGANISM: Artificial Sequence  
 83 <220> FEATURE:

W--> 84 <221> NAME/KEY: p116  
 85 <222> LOCATION: 1..814  
 86 <223> OTHER INFORMATION: corresponds to p116 subunit of eIF3  
 88 <400> SEQUENCE: 4  
 90 Met Gln Asp Ala Glu Asn Val Ala Val Pro Glu Ala Ala Glu Glu Arg  
 91 1 5 10 15  
 94 Ala Glu Pro Gly Gln Gln Gln Pro Ala Ala Glu Pro Pro Pro Ala Glu  
 95 20 25 30  
 98 Gly Leu Leu Arg Pro Ala Gly Pro Gly Ala Pro Glu Ala Ala Gly Thr  
 99 35 40 45  
 102 Glu Ala Ser Ser Glu Glu Val Gly Ile Ala Glu Ala Gly Pro Glu Pro  
 103 50 55 60  
 106 Glu Val Arg Thr Glu Pro Ala Ala Glu Ala Glu Ala Ala Ser Gly Pro  
 107 65 70 75 80  
 110 Ser Glu Ser Pro Ser Pro Pro Ala Ala Glu Glu Leu Pro Gly Ser His  
 111 85 90 95  
 114 Ala Glu Pro Pro Val Pro Ala Gln Gly Glu Ala Pro Gly Glu Gln Ala  
 115 100 105 110  
 118 Arg Asp Glu Arg Ser Asp Ser Arg Ala Gln Ala Val Ser Glu Asp Ala  
 119 115 120 125  
 122 Gly Gly Asn Glu Gly Arg Ala Ala Glu Ala Glu Pro Arg Ala Leu Glu  
 123 130 135 140  
 126 Asn Gly Asp Ala Asp Glu Pro Ser Phe Ser Asp Pro Glu Asp Phe Val  
 127 145 150 155 160  
 130 Asp Asp Val Ser Glu Glu Glu Leu Leu Gly Asp Val Leu Lys Asp Arg  
 131 165 170 175  
 134 Pro Gln Glu Ala Asp Gly Ile Asp Ser Val Ile Val Val Asp Asn Val  
 135 180 185 190  
 138 Pro Gln Val Gly Pro Asp Arg Leu Glu Lys Leu Lys Asn Val Ile His  
 139 195 200 205  
 142 Lys Ile Phe Ser Lys Phe Gly Lys Ile Thr Asn Asp Phe Tyr Pro Glu  
 143 210 215 220  
 146 Glu Asp Gly Lys Thr Lys Gly Tyr Ile Phe Leu Glu Tyr Ala Ser Pro  
 147 225 230 235 240  
 150 Ala His Ala Val Asp Ala Val Lys Asn Ala Asp Gly Tyr Lys Leu Asp  
 151 245 250 255  
 154 Lys Gln His Thr Phe Arg Val Asn Leu Phe Thr Asp Phe Asp Lys Tyr  
 155 260 265 270  
 158 Met Thr Ile Ser Asp Glu Trp Asp Ile Pro Glu Lys Gln Pro Phe Lys

## RAW SEQUENCE LISTING

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:16

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

```

159          275          280          285
162 Asp Leu Gly Asn Leu Arg Tyr Trp Leu Glu Glu Ala Glu Cys Arg Asp
163          290          295          300
166 Gln Tyr Ser Val Ile Phe Glu Ser Gly Asp Arg Thr Ser Ile Phe Trp
167 305          310          315          320
170 Asn Asp Val Lys Asp Pro Val Ser Ile Glu Glu Arg Ala Arg Trp Thr
171          325          330          335
174 Glu Thr Tyr Val Arg Trp Ser Pro Lys Gly Thr Tyr Leu Ala Thr Phe
175          340          345          350
178 His Gln Arg Gly Ile Ala Leu Trp Gly Gly Glu Lys Phe Lys Gln Ile
179          355          360          365
182 Gln Arg Phe Ser His Gln Gly Val Gln Leu Ile Asp Phe Ser Pro Cys
183          370          375          380
186 Glu Arg Tyr Leu Val Thr Phe Ser Pro Leu Met Asp Thr Gln Asp Asp
187 385          390          395          400
190 Pro Gln Ala Ile Ile Ile Trp Asp Ile Leu Thr Gly His Lys Lys Arg
191          405          410          415
194 Gly Phe His Cys Glu Ser Ser Ala His Trp Pro Ile Phe Lys Trp Ser
195          420          425          430
198 His Asp Gly Lys Phe Phe Ala Arg Met Thr Leu Asp Thr Leu Ser Ile
199          435          440          445
202 Tyr Glu Thr Pro Ser Met Gly Leu Leu Asp Lys Lys Ser Leu Lys Ile
203          450          455          460
206 Ser Gly Ile Lys Asp Phe Ser Trp Ser Pro Gly Gly Asn Ile Ile Ala
207 465          470          475          480
210 Phe Trp Val Pro Glu Asp Lys Asp Ile Pro Ala Arg Val Thr Leu Met
211          485          490          495
214 Gln Leu Pro Thr Arg Gln Glu Ile Arg Val Arg Asn Leu Phe Asn Val
215          500          505          510
218 Val Asp Cys Lys Leu His Trp Gln Lys Asn Gly Asp Tyr Leu Cys Val
219          515          520          525
222 Lys Val Asp Arg Thr Pro Lys Gly Thr Gln Gly Val Val Thr Asn Phe
223          530          535          540
226 Glu Ile Phe Arg Met Arg Glu Lys Gln Val Pro Val Asp Val Val Glu
227 545          550          555          560
230 Met Lys Glu Thr Ile Ile Ala Phe Ala Trp Glu Pro Asn Gly Ser Lys
231          565          570          575
234 Phe Ala Val Leu His Gly Glu Ala Pro Arg Ile Ser Val Ser Phe Tyr
235          580          585          590
238 His Val Lys Asn Asn Gly Lys Ile Glu Leu Ile Lys Met Phe Asp Lys
239          595          600          605
242 Gln Gln Ala Asn Thr Ile Phe Trp Ser Pro Gln Gly Gln Phe Val Val
243          610          615          620
246 Leu Ala Gly Leu Arg Ser Met Asn Gly Ala Leu Ala Phe Val Asp Thr
247 625          630          635          640
250 Ser Asp Cys Thr Val Met Asn Ile Ala Glu His Tyr Met Ala Ser Asp
251          645          650          655
254 Val Glu Trp Asp Pro Thr Gly Arg Tyr Val Val Thr Ser Val Ser Trp
255          660          665          670

```

## RAW SEQUENCE LISTING

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:16

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

```

258 Trp Ser His Lys Val Asp Asn Ala Tyr Trp Leu Trp Thr Phe Gln Gly
259      675      680      685
262 Arg Leu Leu Gln Lys Asn Asn Lys Asp Arg Phe Cys Gln Leu Leu Trp
263      690      695      700
266 Arg Pro Arg Pro Pro Thr Leu Leu Ser Gln Glu Gln Ile Lys Gln Ile
267 705      710      715      720
270 Lys Lys Asp Leu Lys Lys Tyr Ser Lys Ile Phe Glu Gln Lys Asp Arg
271      725      730      735
274 Leu Ser Gln Ser Lys Ala Ser Lys Glu Leu Val Glu Arg Arg Arg Thr
275      740      745      750
278 Met Met Glu Asp Phe Arg Lys Tyr Arg Lys Met Ala Gln Glu Leu Tyr
279      755      760      765
282 Met Glu Gln Lys Asn Glu Arg Leu Glu Leu Arg Gly Gly Val Asp Thr
283      770      775      780
286 Asp Glu Leu Asp Ser Asn Val Asp Asp Trp Glu Glu Glu Thr Ile Glu
287 785      790      795      800
290 Phe Phe Val Thr Glu Glu Ile Ile Pro Leu Gly Asn Gln Glu
291      805      810

```

294 &lt;210&gt; SEQ ID NO: 5

295 &lt;211&gt; LENGTH: 106

296 &lt;212&gt; TYPE: PRT

297 &lt;213&gt; ORGANISM: Artificial Sequence

299 &lt;220&gt; FEATURE:

W--&gt; 300 &lt;221&gt; NAME/KEY: p116

301 &lt;222&gt; LOCATION: 175..279

302 &lt;223&gt; OTHER INFORMATION: corresponds to a portion (RRM) of eIF3 p116 subunit

304 &lt;400&gt; SEQUENCE: 5

306 Met Asp Arg Pro Gln Glu Ala Asp Gly Ile Asp Ser Val Ile Val Val

307 1 5 10 15

310 Asp Asn Val Pro Gln Val Gly Pro Asp Arg Leu Glu Lys Leu Lys Asn

311 20 25 30

314 Val Ile His Lys Ile Phe Ser Lys Phe Gly Lys Ile Thr Asn Asp Phe

315 35 40 45

318 Tyr Pro Glu Glu Asp Gly Lys Thr Lys Gly Tyr Ile Phe Leu Glu Tyr

319 50 55 60

322 Ala Ser Pro Ala His Ala Val Asp Ala Val Lys Asn Ala Asp Gly Tyr

323 65 70 75 80

326 Lys Leu Asp Lys Gln His Thr Phe Arg Val Asn Leu Phe Thr Asp Phe

327 85 90 95

330 Asp Lys Tyr Met Thr Ile Ser Asp Glu Trp

331 100 105

334 &lt;210&gt; SEQ ID NO: 6

335 &lt;211&gt; LENGTH: 33

336 &lt;212&gt; TYPE: DNA

337 &lt;213&gt; ORGANISM: Artificial Sequence

339 &lt;220&gt; FEATURE:

340 &lt;221&gt; NAME/KEY: primer\_bind

341 &lt;222&gt; LOCATION: 1..33

342 &lt;223&gt; OTHER INFORMATION: HCV RRM 5' primer (RRMfwd)

## RAW SEQUENCE LISTING

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:16

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

```

344 <400> SEQUENCE: 6
346 catatggatc ggccccagga agcagatgga atc 33
349 <210> SEQ ID NO: 7
350 <211> LENGTH: 33
351 <212> TYPE: DNA
352 <213> ORGANISM: Artificial Sequence
354 <220> FEATURE:
355 <221> NAME/KEY: primer_bind
356 <222> LOCATION: 1..33
357 <223> OTHER INFORMATION: HCV RRM 3' primer (RRMrev)
359 <400> SEQUENCE: 7
361 gtgctcgagc cactcgtcac tgategtcat ata 33
364 <210> SEQ ID NO: 8
365 <211> LENGTH: 29
366 <212> TYPE: DNA
367 <213> ORGANISM: Artificial Sequence
369 <220> FEATURE:
370 <221> NAME/KEY: primer_bind
371 <222> LOCATION: 1..29
372 <223> OTHER INFORMATION: HCV IRES 5' primer (IRESfwd)
374 <400> SEQUENCE: 8
376 accgctagcc tcccctgtga ggaactact 29
379 <210> SEQ ID NO: 9
380 <211> LENGTH: 46
381 <212> TYPE: DNA
382 <213> ORGANISM: Artificial Sequence
384 <220> FEATURE:
385 <221> NAME/KEY: primer_bind
386 <222> LOCATION: 1..46
387 <223> OTHER INFORMATION: HCV IRES 3' primer (IRESrev)
389 <400> SEQUENCE: 9
391 gaaagctttt ttctttgagg tttaggattt gtgctcatga tgcacg 46
394 <210> SEQ ID NO: 10
395 <211> LENGTH: 95
396 <212> TYPE: DNA
397 <213> ORGANISM: Artificial Sequence
399 <220> FEATURE:
400 <221> NAME/KEY: primer_bind
401 <222> LOCATION: 1..95
402 <223> OTHER INFORMATION: primer IIIabcfwd which corresponds to T7 polymerase promoter
+ 139-215 of
403 HCV (regions IIIa-IIIb)
405 <400> SEQUENCE: 10
407 taatacgact cactataggg tagtggtctg cggaaccggt gactacaccg gaattgccag 60
409 gacgaccggg tcctttcttg gataaaccgg ctcaa 95
412 <210> SEQ ID NO: 11
413 <211> LENGTH: 60
414 <212> TYPE: DNA
415 <213> ORGANISM: Artificial Sequence
417 <220> FEATURE:

```



RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:17

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:10; Line(s) 402

**VERIFICATION SUMMARY**

DATE: 06/28/2005

PATENT APPLICATION: US/10/538,471

TIME: 08:40:17

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\06282005\J538471.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:27 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1  
L:52 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:2  
L:69 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3  
L:84 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:4  
L:300 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5